

UNITED STATES DISTRICT COURT
DISTRICT OF MINNESOTA

IN RE: BAIR HUGGER FORCED AIR
WARMING PRODUCTS LIABILITY
LITIGATION

This Document Relates to All Actions

MDL No. 15-2666
(JNE/FLN)

**PRETRIAL ORDER NO. 12:
Computer Assisted Review Protocol**

PROTOCOL REGARDING COMPUTER ASSISTED REVIEW

Pursuant to the Parties' agreement, the procedures outlined herein shall govern the use of computer assisted review ("CAR") for electronically stored information ("ESI") during the pendency of this litigation.

I. Data Sources. To the extent either party collects e-mail or other electronic documents, they will do so for at least the data sources agreed upon with the opposing party. Non-custodial data sources, such as relevant databases or group shares, shall also be discussed. The Parties agree that the stipulations set forth in this Protocol shall not limit their ability to request discrete documents or document categories that may not belong to a particular agreed-upon custodian or data source. Nothing in this Protocol shall be construed to waive the Parties' rights to request documents or to object to particular discovery requests as they are entitled under the Federal Rules of Civil Procedure. The Parties will confer in good faith upon any such requests or objections. If the Parties cannot reach agreement, any dispute shall be presented to the Court in accordance with the Federal Rules of Civil Procedures, Local Rules, and orders of the Court.

II. Computer Assisted Review. The parties reserve the right to use CAR, also referred to "technology assisted review," or "predictive coding," in which a machine-

learning algorithm assesses the likely relevance of a corpus of documents based on manual relevance determinations applied to a subset of that corpus. Although the Parties presume that each party is in the best position to know how to search for its own responsive ESI that will be considered for CAR, the Parties have the right to make reasonable inquiries concerning the adequacy and success of the other's process. As with key word search term negotiation, the use of CAR is best when both parties can agree as to what documents constitute responsive versus non-responsive material. The goal here is to get to the facts and not be bogged down with technological minutia. The following protocol is designed so as not to be unduly burdensome, while also allowing both Parties to reap the benefits of using CAR, including greater accuracy than keyword search or manual review, reduced costs, and the more rapid document delivery time frames offered by CAR. To the extent the Protocol does not achieve these goals in practice, the Parties reserve their rights to meet-and-confer and seek appropriate modification. The protocol for the review and production of documents using CAR is as follows:

a. **Documentation of the CAR Startup Process.** Each party leveraging CAR will produce a written description, with reasonable levels of detail, of the method or methods used to search for documents and data responsive to the properly served requests for production, that will be considered for CAR, including but not limited to the custodial and non-custodial sources being searched, the rationale for that selection, the specific software being used for CAR, and the specifications being implemented.

Once this has been negotiated and agreed to as being satisfactory, CAR seed set training will commence.

b. **Training Methodology.** At producing party's sole option, but subject to receiving party's right to raise this issue with the Court per paragraph II.g.4.iv if the producing party elects not to follow this procedure, the following procedure will be used to designate individuals who will be responsible for document assessment and generation of the CAR document training set. Each party will designate two subject-matter experts. These designated persons will work together at the same time, in the same room, to determine which documents are either responsive or non-responsive. These designated persons are not allowed any form of phone, computer, tablet, and paper or information recording device. What is seen in the room, stays in the room. Furthermore, these experts shall agree in writing not to disclose the content of nonresponsive or privileged documents to anyone other than the producing party and will be subject to sanction for any disclosure or improper use. Whether or not the producing party elects this procedure, the remainder of this protocol shall be used by the producing party in the performance of CAR.

(1) **Initial CAR Training:** Building the initial training model for the computer can be a time consuming process, while subsequent training across newly added documents takes less effort. For the producing party's initial model build: the Parties may each identify up to 250 unique production documents to be included in the initial

feed into the system to speed up training. In addition to the production documents above, no less than 1500 randomly selected documents will be reviewed and categorized by the producing party for initial training.

c. **Privileged Documents.** Documents that are privileged are often good training materials for the CAR system. If a document, at first glance, is determined to be privileged, the experts from the opposing party, if they are in the room, are to look away, while the privilege content is assessed solely by the producing party. If this content contains fact patterns outside of this matter, the document should be excluded, otherwise if relevant, it should be included. Post-CAR-training, these privileged documents can be filtered and removed from the final production. Training the system as outlined above using privileged documents does not result in any disclosure of the privileged document to the opposing party and does not constitute a waiver of privilege.

d. **Documents Not Suitable for CAR.** Documents that are not appropriate for CAR retrieval, including but not limited to spreadsheets, databases, flat-file PDF documents, Computer-Aided-Design files, video, audio, and images, will be not be subject to the CAR process, but may still be requested to be produced in response to properly served requests for production.

e. **Training Targets.** Once adequate training of the system has been achieved, the CAR process will categorize the target documents. Adequate training of the CAR system shall be reached when both Recall rate is at or exceeds 80% and the

producing party determines in good faith that the burden of additional enhancement of Recall Rate outweighs the benefit. If the producing party is unable to meet the 80% threshold for Recall with reasonable effort, the Parties shall meet and confer on other steps that can reasonably be taken to meet that threshold and/or modification of the threshold if necessary. If the Parties cannot agree on other steps and/or modification of the threshold, the Parties shall bring the issue to the Court for resolution.

f. **Use of Exemplars and Remediation.** The Parties can also elect to use certain exemplars of good or bad documents obtained through keyword searching or other means to adjust the training of the CAR system so that its overall accuracy is improved at any time during the process. However, keyword search, concept search, or other similar methodologies will not be used prior to predictive coding being run, unless and until, for a specific collection of documents, the rate of prevalence is too low to achieve the targeted level of performance via normal means. This procedure contains a number of remediation steps that can be taken prior to the need to perform such actions and these are to be followed until such time as to be deemed ineffective. If the producing party believes keyword search, concept search or other similar methodologies has become necessary, the Parties will meet and confer regarding the use of such methodology, and if no agreement can be reached, the issue shall be brought to the Court for resolution.

g. **Measurements and Documentation.** The following pre- and post-run information will be captured, documented, and shared for each and every run in order

to determine that the CAR training was appropriately carried out and in order to determine that Recall rate in II.e is being met:

- (1) A written description of the methods that were used to assess the adequacy of the process used.
- (2) The following assessment information:
 - i. The number of documents (or other ESI units) that were collected by the producing party and the means used for identifying those documents.
 - ii. The steps that were taken to evaluate the success of any method, other than the CAR contemplated herein, used to reduce the number of documents collected to hold down the number that was ultimately submitted to the CAR system (culling evaluation).
 - iii. The number and percentage of total documents submitted to the CAR system as the post-culling population.
 - iv. The number and percentage of the documents that were ultimately deemed responsive, exclusive of their non-responsive family members. That is, the number of documents that were actually identified by the CAR system as responsive.
 - v. The levels of Recall and Precision that were achieved, and how these were calculated. These will be reported both pre-run and post-run, but prior to audit.

- (3) **Audit.** The results of each predictive coding run will be audited using the following process:
- i. A random sample of 1,000 documents drawn from the set of documents, post-categorization, that were not considered potentially responsive and therefore not produced (called the “negative set”).
 - ii. A similar random sample will also be produced from the responsive set.
 - iii. The method used to select the random documents will be described by producing party detailing the technical methodology and/or workflow used to create the sample document sets.
 - iv. Both sets of randomly selected documents will be audited for accuracy, relative to the set of documents they represent. The persons who will perform the audit shall be the same individuals who conducted the initial training.
 - v. Prior to this audit occurring, the documents will be rapidly assessed for privilege solely by the producing party and those privileged documents withheld from confirmatory review. These privileged documents will still be included in the overall accuracy count for purposes of assessment. The objective here is

to determine the effectiveness of the CAR training model and its overall performance.

- vi. The goal of the audit is to assess if potentially responsive materials were located in the negative set, and if responsive materials contain content that should be in the negative set. The goal of the measurement is to determine the overall frequency of such wrong designation, and to assess whether this is acceptable or otherwise based upon agreed to performance metrics and thresholds above.
- vii. Up to 400 document moves can occur per audit (irrelevant to relevant and vice-versa). Once these changes have occurred, predictive coding can be re-run and accuracy re-assessed.
- viii. The number of documents moved and in which direction they moved will be documented.
- ix. The performance ratios obtained through this measurement process will be documented, capturing the effective precision and recall rates, as well as the number of privileged documents that were withheld from confirmatory review.
- x. If the recall rate agreed to above is not obtained, the process will iterate until it is.

- xi. For the initial set of documents used for training, and for each pass through this QC process promptly after each pass being completed, regardless of whether the targeted recall rate is obtained, not obtained, or exceeded, the producing party will make available to the receiving party a log in Excel format containing:
- a. the CAR run date & time;
 - b. a list of control numbers for all documents the system identified as responsive, along with a copy of those documents except that no copy need be provided for those documents withheld for privilege;
 - c. a list of control numbers for all documents the system identified as non-responsive, along with a copy of those documents the system incorrectly identified as non-responsive and which were moved from the non-responsive set to the responsive set, except that no copy need be provided for those documents withheld for privilege;
 - d. a count of the number of privileged documents correctly identified as responsive but withheld for privilege;
 - e. a count of the number of privileged documents incorrectly identified as responsive but withheld for privilege;

- f. the metadata for all documents as described in Pretrial Order No. 10 at paragraph II.P or Pretrial No. 11 depending on the privilege status of the document;
- g. four additional columns in the spreadsheet indicating: (1) whether the document was marked by the CAR system as responsive or non-responsive; (2) whether, after manual review, the document was changed from the non-responsive set to the responsive or vice-versa; (3) if a copy of the document is not produced by reason of privilege, the reason for withholding the document on the basis of privilege; and 4) for any document for which a copy is not being produced due to non-responsiveness or privilege, information sufficient to allow the receiving party to determine why the document was considered non-responsive and/or privileged.¹

- xii. The receiving party has fourteen (14) calendar days to review the log and associated production and meet and confer to request or suggest changes and, if no agreement can be reached, bring the issue to the Court for resolution. The log shall be treated as

¹ Providing the level of detail required for privilege logs by Pretrial Order No. 11 satisfies this provision for privileged documents.

Confidential under the Protective Order and shall not be used for any purpose other than as provided by this Protocol.

- xiii. If during the audit it is found that a certain discrete class or sub-grouping of documents are being incorrectly selected by the predictive coding system, it is permissible to filter for these and remove these from the initial document population prior to predictive coding be re-run. If the producing party believes this is necessary, however, the Parties shall first meet & confer as to the reasons for the filter becoming necessary and the filtering process, and if no agreement can be reached, bring the issue to the Court for resolution.

(4) **Audit Assessment.** Based on the receiving party's evaluation of the information the providing party has produced under the paragraphs 2 & 3 above and for each production of documents, the receiving party may, at its election, do one or more of the following:

- i. Accept the use of CAR as is;
- ii. Request a meet-and-confer to discuss additional custodians, time ranges, key words or other means for identifying new documents to be considered for responsiveness;
- iii. Request a meet-and-confer to discuss modifications to the CAR process; and/or

- iv. Ask for Court intervention if the Parties cannot agree to an appropriate resolution through steps i-iii above, including but not limited to asking the Court to allow the receiving party to participate in the training under paragraph II.b and audit of the CAR system under paragraph II.g.3.
- (5) The Parties will attempt in good faith to resolve any disagreements through a meet and confer process. Should the Parties be unable to reach agreement, the dispute will be submitted to the Court for resolution.
- (6) CAR is an iterative process. As new document sources are selected for review or additional information is obtained by the Parties, additional training and ongoing assessment of the CAR system will be performed using the protocols set out above.

IT IS SO ORDERED.

Date: July 8, 2016

s/ Joan N. Ericksen
JOAN N. ERICKSEN
United States District Judge